

ROBUST PARKING SPACE DETECTION CONSIDERING INTER-SPACE CORRELATION

Qi Wu¹, Chingchun Huang², Shih-yu Wang³, Wei-chen Chiu³, Tsuhan Chen¹

Electrical & Computer Engineering Department, Carnegie Mellon University¹
Industrial Technology Research Institute²
National Chiao Tung University³

ABSTRACT

A major problem in metropolitan areas is searching for parking spaces. In this paper, we propose a novel method for parking space detection. Given input video captured by a camera, we can distinguish the empty spaces from the occupied spaces by using an 8-class Support Vector Machine (SVM) classifier with probabilistic outputs. Considering the inter-space correlation, the outputs of the SVM classifier are fused together using a Markov Random Field (MRF) framework. The result is much improved detection performance, even when there are significant occlusion and shadowing effects in the scene. Experimental results are given to show the robustness of the proposed approach.

1. INTRODUCTION

In recent years, parking has become a serious problem with the increase of private vehicles. Looking for parking spaces always wastes travel time. For the driver's convenience, public parking lots should provide the location of available parking spaces. However, maintaining such information manually needs lots of human resource. Therefore, automatic parking space detection has been employed in many systems for counting the number of available parking spaces, identifying their locations and monitoring changes of their status over time.

Toward this need, many researchers have been improving parking space detection systems. Foresti et al.[1] used visual surveillance, which requires real-time interpretation of image sequences. Wang and Hanson [2] extracted and analyzed the structured geometric information of parking lots from aerial images. Lee et al.[3] and Masaki [5] kept tracking and recording the movement of vehicles for finding the empty parking space. Gupte et al.[4] calculated the width and length of vehicles based on the relations of geometrical shapes and triangle function. These detection methods require high computation and large storage. In this paper, we present a novel method for parking space detection using only a few frames captured from a single camera. Considering the position of camera which is impossible to be set high above the parking lot and the certain correlation between neighboring spaces,

our detection method is robust to the influence of light variability, shadow and occlusion. To obtain high detection accuracy under these critical conditions, we train and recognize empty parking spaces by applying machine learning methods instead of directly segmenting the vehicle out of each parking space. Our goal is to build a highly accurate automatic detection system which is stable for real-time applications.

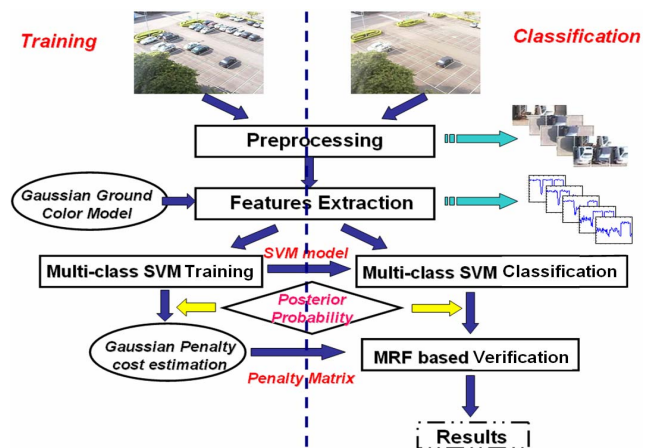


Fig. 1. System overview

2. SYSTEM OVERVIEW

The outline of the proposed parking space detection system is shown in Fig.1. This system consists of four parts: preprocessing, ground model feature extraction, multi-class SVM recognition and MRF based correction. First, we preprocess the input frames and divide them into small patches which contain 3 parking spaces each. Then, a gaussian ground model is set up to obtain the likelihood of ground for the pixels in the patches as our features. Next, multi-class Support Vector Machine (SVM) is trained to analyze and classify the patches into 8 classes of parking space status. Finally, Markov Random Field (MRF) are builded to solve the conflicts between two neighboring patches in order to improve the recognition accuracy.

3. PREPROCESSING

Given an input video frame (Fig.2(a)), the parking regions can be easily obtained as shown in Fig.2(b), assuming that we know the intrinsic and extrinsic parameters of the camera. In practice, due to vehicle shadow and occlusion, if a basic single parking space is chosen for detection, the result is not good enough. Thus, in order to use the inter-space correlation, 3 parking spaces are proposed as a detection patch, which contains the space under consideration and two neighboring spaces. Using perspective transformation, the original patches can be normalized into rectangular ones (Fig.2(c)). Additionally, in the training process, we classify them into 8 (2^3) statuses as in Fig.2(d) while labelling an empty space as 0 and an occupied space as 1.

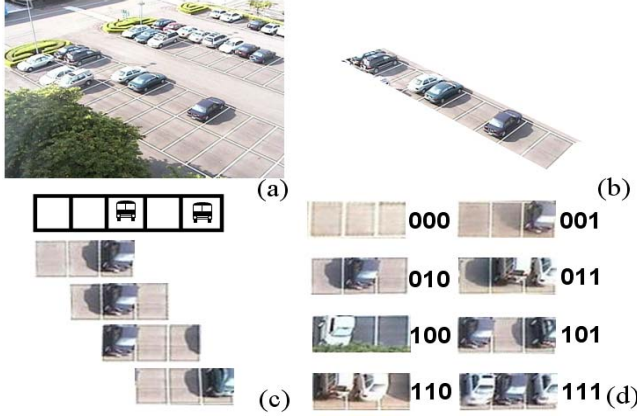


Fig. 2. Preprocess the input frame and generate the patches. (a) origin video frame (b) one detection row (c) patch generation (d) 8 space statuses

4. FEATURE EXTRACTION

In order to segment the car regions from the ground regions, we need a reliable ground color model that adapts to different ground colors and different lighting conditions. Color histogram analysis reveals that the distribution of ground colors is clustered in the chromatic color space and this distribution can be represented by a gaussian model. Therefore, we can obtain the likelihood of ground for any pixel x as follows

$$L(x) = \exp\left(-\frac{1}{2}(x - m_g)\Sigma_g^{-1}(x - m_g)^t\right) \quad (1)$$

where m_g is the mean and Σ_g is the variance of ground color distribution.

Three scanning lines are used to extract the color of pixels in the patch from left to right. We computed the likelihood of each pixel using the ground color model. For a single patch, since there are 75 pixels along each scanning line, 215 (75×3 scanlines) features are extracted from one patch. Moreover,

Principle Component Analysis (PCA) is applied to extract the critical features. After dimension reduction, 50 critical features are picked out for further training, which contain over 99% of the complete energy.

5. RECOGNITION

Given the features extracted from each patch, Support Vector Machine (SVM) [6] is applied in order to identify empty space recognition. SVM developed by Vapnik is a popular binary classifier in practical application. It maps vectors with appropriate kernel function into high-dimension of the feature space and satisfy the linear separable constraint:

$$\min |w^T x_i + b| = 1, i = 1, \dots, N \quad (2)$$

Unlike the classical SVM which uses Signum function only for binary output (+1 or -1), we need to know the posterior probability of every status. Thus, binary posterior probability is obtained as follows:

$$p(y_i = \pm 1 | x, w) = \frac{1}{1 + \exp(-1/||w||f(x))} \quad (3)$$

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x, x_i) \quad (4)$$

where α_i denotes the Lagrange multipliers, $\{(x_i, y_i) | x_i \in R, y_i = \pm 1, i = 1, \dots, N\}$ denotes a set of training samples, $K(x, x_i)$ is the kernel function, and $1/||w||$ is the distance between the hyperplane (w, b) and the support vectors.

Moreover, we adapt the general binary SVM classifier for multi-class problem by using one-against-one strategy, which takes all possible two-class combinations. Therefore, $N(N-1)/2$ SVMs are trained and each SVM classifier separates a pair of classes. Here, N is the number of classes. Finally, Radial Basis Function (RBF) is adopted to be the kernel function of the classifier.

Trained by the $8 \times (8-1)/2 = 28$ binary classifiers, given features x , the probability of the patch belong to the i th status can be represent as:

$$p(y_i | x) = \frac{1}{2 - N + \sum_{j=1, i \neq j}^N \frac{1}{p_{ij}(y_i | x)}}, i = 1, \dots, 8 \quad (5)$$

where $p_{ij}(y_i | x)$ is binary posterior probability obtained by Equation (3). Knowing these probabilities, we can say a patch is in the i th status if $p(y_i | x) = \max_{j=1, \dots, N} p(y_j | x)$, where $i \in j, N = 8$

6. OPTIMIZATION AND CONFLICT CORRECTION

Same as any other machine learning algorithms, SVM cannot guarantee perfect performance in classification. In the parking space detection, because each pair of neighboring patches

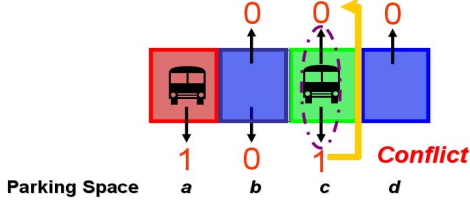


Fig. 3. Conflict between two neighboring patches. The result of SVM shows that a is the space occupied by vehicle, b and d are the empty spaces while c is the space in conflict.

have 2 shared parking spaces, conflict may occur when one or both of them are classified into wrong statuses as shown in Fig.3, we can not ensure whether the parking space c is occupied or not just depending on the results of SVM. Therefore, Markov Random Field (MRF) [7] is applied for optimizing the results of SVM to correct these conflicts.

6.1. Markov Random Field Based Correction

In our MRF framework as shown in Fig.4, we define a patch labelling problem as assigning every patch k in one parking row a label n . These statuses labelled $S_n^{(k)}$ are independent and identically distributed when the posterior probability $X^{(k)}$ of SVM is given. So the log-likelihood function can be presented as:

$$\begin{aligned}
 l(S) &= \log \prod_K p(S^{(k+1)}, S^{(k)}) \\
 &= \log \prod_K \prod_N p(S_n^{(k)} | X^{(k)}) p(S_n^{(k+1)} | S_n^{(k)}) \\
 &= \sum_K \sum_N \log p(S_n^{(k)} | X^{(k)}) \\
 &\quad + \sum_K \sum_N \log p(S_n^{(k+1)} | S_n^{(k)})
 \end{aligned} \tag{6}$$

where K is the number of patches in one detection row and $N = 8$ is the number of statuses.

Hence, the MRF energy function E , which can be viewed as the log likelihood of the posterior distribution of SVM, is composed by the data energy E_d and the smoothness energy E_s . The data energy is the sum of per-patch data cost $d_k(S_n^{(k)})$, which equals to the negative log posterior probability of SVM result: $-\log p(S_n^{(k)} | X^{(k)})$, that $E_d = \sum_K d_k(S_n^{(k)})$. The smoothness energy is defined as the sum of horizontal neighboring penalty cost, $V_k(n, m) = -\log p(S_n^{(k+1)} | S_m^{(k)})$, that $E_s = \sum_K V_k(n, m) \quad m, n \in N$. Hence, in order to solve the energy minimization problem, our key task is how to train and estimate the appropriate penalty cost.

6.2. Penalty Cost Estimation

Since there are 2 overlapping parking spaces between pairs of neighboring patches, three kinds of relationship can be appeared between them.

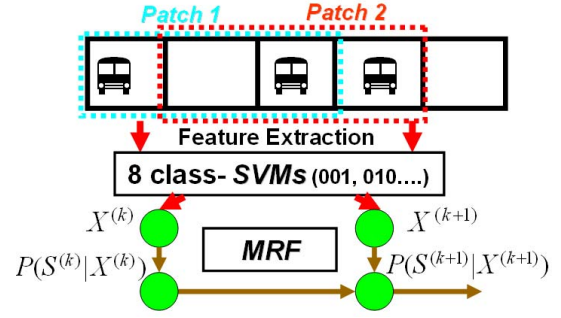


Fig. 4. Markov random field based correction

- **No conflict** e.g. (100 & 000), (101 & 011)...
- **One space in conflict** e.g. (000 & 100), (111 & 011)...
- **Two spaces in conflict** e.g. (110 & 010), (001 & 101)...

Assume n, m are the labels of neighboring patches observed from the results of SVM, n', m' are their ground truth, the penalty cost of pair (n, m) can be trained as:

$$V(n, m) = \frac{\sum_{k \in T_{(n,m)}} (||d_k(S_n) - d_k(S_{n'})|| + ||d_k(S_m) - d_k(S_{m'})||)}{N_{T_{(n,m)}}} \tag{7}$$

Here, $T_{(n,m)}$ is training set of all neighboring patches observed label (n, m) from SVM while $N_{T_{(n,m)}}$ is its size. Defined 8 classes of status, there are $8 \times 8 = 64$ penalty costs estimated as shown in Tab. 1. Using this pre-computed penalty matrix and proposed MRF framework, we can easily solve the conflicts and improve detection performance.

	000	001	010	011	100	101	110	111
000	0.00	1.18	1.32	2.10	0.00	1.31	1.41	2.22
001	0.00	1.21	1.28	1.90	0.00	1.25	1.32	1.97
010	1.36	0.00	1.87	1.17	1.02	0.00	1.88	1.52
011	1.28	0.00	1.98	1.22	0.99	0.00	1.95	1.61
100	1.57	2.13	0.00	1.31	1.13	2.05	0.00	1.32
101	1.56	2.05	0.00	1.27	1.20	1.93	0.00	1.44
110	1.92	1.43	1.37	0.00	2.21	1.42	1.28	0.00
111	2.03	1.52	1.34	0.00	2.11	1.47	1.26	0.00

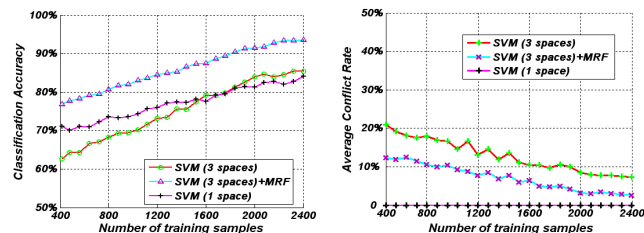
Table 1. The matrix of penalty cost between neighboring patches. The first row represents the statuses of the former patch while the first column represents the statuses of the later patch.

7. EXPERIMENTS AND RESULTS

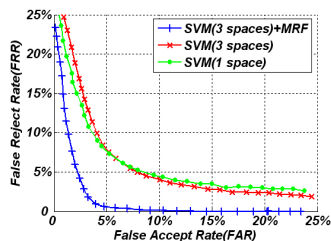
We captured the video frames from the real scene. First, A total of 300 ground samples from different ground colors and

lighting conditions were used to determine the color distribution of ground in chromatic color space. Then, we generated 2400 patches (300 for each class of status) from 500 frames as the training data. Finally, after obtaining gaussian ground model and training 8-class SVM classifier, we evaluated the performances of our algorithm using 300 test frames. Moreover, we also compare the results before and after MRF conflict correction, to prove the the performance improvement.

With the increase of training samples, Fig.5(a) demonstrates the rate of successfully distinguishing the available parking spaces, while Fig.5(b) demonstrates rate of conflict between pairs of neighboring patches. Fig.5(c) shows evaluation of the False Accept Rate (FAR) and False Reject Rate (FRR) performance. Here, FAR is defined as the rate that occupied spaces are misclassified into empty spaces, while FRR is defined as the rate that available parking spaces are misclassified into spaces parked with vehicles.



(a) Classification Accuracy (b) Average Conflict Rate



(c) Evaluation of the FAR and FRR performance

Fig. 5. Comparison of Experiment Results between single space SVM detection, 3 spaces SVM detection and its MRF correction

From these plots, it can be pointed out that the classification accuracy increased with more training samples. Thus, less conflicts will occur. In addition, it is obvious from the figure that comparing with SVM using single space and 3 spaces for detection, which has 84.35% and 85.57% in accuracy, Markov Random Field based correction can improve the precise to 93.52% and sharply reduce the average conflict rate from 7.32% to 2.57%. Moreover, from the evaluation of FAR and FRR in Fig.5(c), we can note that with the same FRR, FAR of SVM+MRF is lower, while with the same FAR, its FRR is lower. The FAR and FRR under optimal performance are shown in the Tab.2, which indicate the obvious improvement of the MRF correction. It can be attributed to the fact

that the posterior probabilities of truth between neighboring patches are always close to those of wrong results given by SVM, which can be easily corrected using MRF.

	FAR	FRR
SVM(1 space)	4.85%	8.12%
SVM(3 spaces)	4.39%	8.73%
SVM(3 spaces)+MRF	1.25%	3.56%

Table 2. The FAR and FRR under optimal performance

8. CONCLUSION

In this paper, we proposed a novel method for parking space detection. After preprocessing the input image into patches of 3 spaces each, multi-SVM with probabilistic outputs is applied to these patches to recognize spaces and describe the relationship of neighboring patches. Finally, by applying Markov Random Field to solve the potential conflict, we can optimize the result and improve the recognition accuracy. It can be seen from the experiments that our method is robust and effective.

9. REFERENCES

- [1] G. L. Foresti, C. Micheloni and L. Snidaro, "Event classification for automatic visual-based surveillance of parking lots," *Proc. of the 17th International Conference on Pattern Recognition*, Vol.3, pp.314-317, 2004
- [2] X. G. Wang, A. R. Hanson "Parking lot analysis and visualization from aerial images," *Proc. in Fourth IEEE Workshop Applications of Computer Vision*, pp.36-41,1998
- [3] C. H. Lee, M. G. Wen, C. C. Han and D. C. Kou, "An automatic monitoring approach for unsupervised parking lots in outdoors," *Security Technology, 2005. CCST '05. 39th Annual 2005 International Carnahan Conference*, pp.271 - 274, 2005
- [4] S. Gupte, O. Masoud and P. Papanikolopoulos, "Vision-based vehicle classification," *IEEE Conference on Intelligent Transportation System*, pp.46-51, 2000
- [5] I. Masaki, "Machine-vision systems for intelligent transportation systems," *IEEE Conference on Intelligent Transportation System*, Vol.13(6), pp.24-31, 1998
- [6] J. Platt, "Probabilistic outputs for support vector machines and comparison to regularized likelihood methods." *Advances in Large Margin Classifiers*, Cambridge, MA, 2000. MIT Press.
- [7] Y. Boykov, O. Veksler and R. Zabih, "Efficient Approximate Energy Minimization via Graph Cuts" *IEEE transactions on PAMI*, Vol.3, pp.1222-1239, 2001